

## BAB III METODOLOGI

### 3.1 Pengumpulan Data

Dalam penelitian ini, menurut cara mendapatkan data yang digunakan dapat dikategorikan sebagai data sekunder. Data sekunder adalah data yang tidak diperoleh langsung oleh peneliti, data di sini bisa dalam bentuk dokumen atau arsip yang dimiliki oleh lembaga atau seseorang yang menjadi subjek penelitian (SPSS Statistik, 2018). Data yang digunakan pada penelitian merupakan arsip dari penyedia citra satelit komersial yang disebut Planet. Lembaga ini menggunakan beberapa satelit kecil untuk mengambil gambar seluruh Bumi setiap hari. Sumber daya peralatan yang terbatas mempersulit peneliti untuk mendapatkan data citra satelit secara langsung (data primer), sehingga data sekunder digunakan.

Planet mengoperasikan 120+ Dove, 15 SkySat dan 5 satelit RapidEye seperti pada Gambar 2.1 yang menyediakan dataset serbaguna untuk analisis geospasial pasar, lingkungan dan perubahan global. Planet memberikan data pencitraan Bumi pada skala yang belum pernah disediakan perusahaan lain sebelumnya (PLANET.COM, 2019).

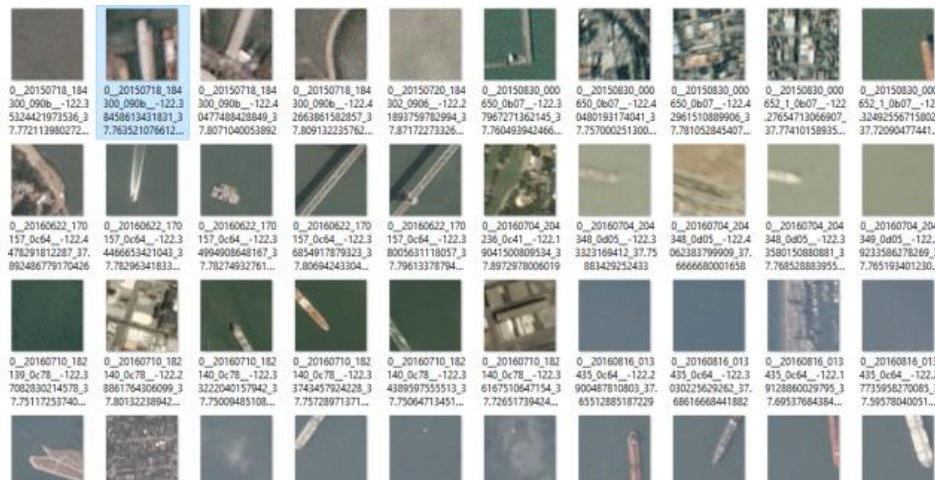
*Training* set atau data latih yang digunakan adalah citra satelit berjenis sensor optik seperti pada yang dikelompokkan menjadi 1000 citra “kapal” dan 3000 citra “bukan kapal”. Berikut keterangan terkait dengan data *training* yang akan digunakan:

- a. Citra satelit untuk wilayah Teluk San Francisco dan Teluk San Pedro di California, Amerika Serikat.
- b. Citra satelit yang digunakan adalah capture penampakan kawasan teluk dari atas.
- c. Jenis citra satelit yang digunakan merupakan citra sensor optik atau tidak dapat menembus awan.
- d. Citra yang digunakan adalah citra berwarna RGB.
- e. Data terdiri atas 4000 citra dengan ukuran 80x80 pixel.
- f. Data terdiri dari 1000 citra “kapal” dan 3000 citra “non kapal”.
- g. Data *training* yang diberi label sebagai kapal adalah citra dengan penampakan kapal secara utuh.
- h. Data *training* yang diberi label sebagai non-kapal merupakan penampakan citra bukan kapal dan kapal yang tidak utuh (hanya sebagian yang terlihat).

- i. File berformat JSON yang berisi data, label, nomor adegan, dan metadata lokasi.
- j. Format citra adalah PNG atau (.png).



Gambar 3. 1. Data set dengan label 1 (*ship*)

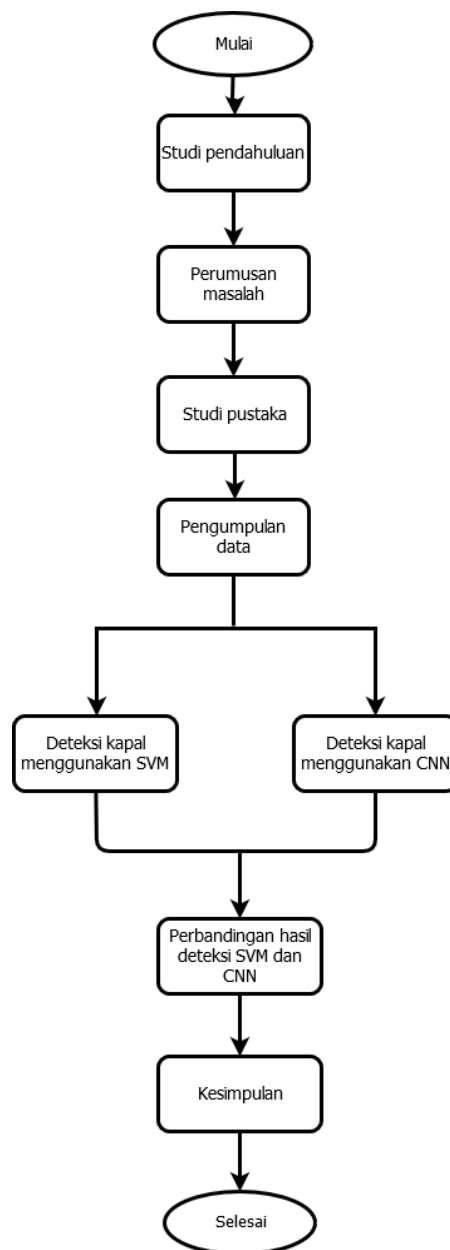


Gambar 3. 2. Data set dengan label 0 (*non-ship*)

## 3.2 Perancangan

### 3.2.1 Perancangan Penelitian

Perancangan adalah gambaran rancangan penelitian secara umum dan terstruktur, serta memudahkan peneliti dalam implementasi program. Pada penelitian ini perancangan dibuat dalam bentuk *flowchart* atau diagram alir seperti pada Gambar 3. 3 yang menggambarkan bagaimana penelitian ini akan dikembangkan mulai dari studi pendahuluan sampai dengan kesimpulan.



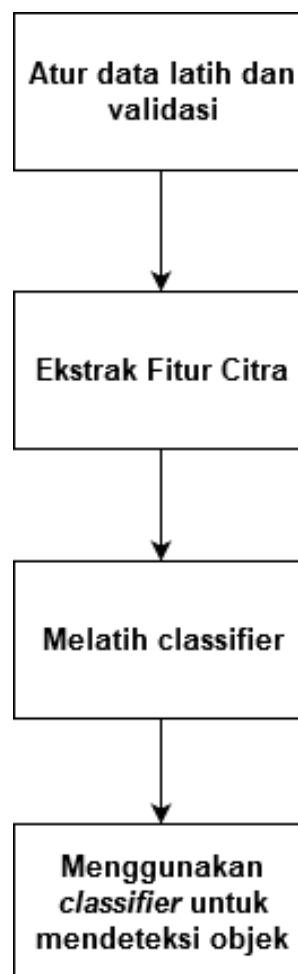
Gambar 3. 3 *Flowchart* penelitian

Penjelasan *flowchart* penelitian:

- a. Studi pendahuluan dilakukan untuk memperoleh masukan mengenai objek yang akan diteliti. Melalui proses studi ini, dapat diperoleh informasi terkait permasalahan yang diangkat dalam penelitian dan variabel-variabel lain terkait dengan penelitian ini.
- b. Pada tahap perumusan masalah dilakukan penentuan permasalahan yang dihadapi oleh dalam penelitian, sehingga dapat dirumuskan masalah dalam bentuk pertanyaan.
- c. Studi pustaka dilakukan dengan mencari dan mempelajari berbagai referensi yang mengacu kepada suatu sumber sumber, baik dari penelitian-penelitian sebelumnya, buku dan sumber lain yang dijadikan referensi untuk memperoleh data dan teori-teori yang dibutuhkan untuk mendukung dilakukannya penelitian.
- d. Pengumpulan data dilakukan untuk mengumpulkan data yang diperlukan untuk penelitian.
- e. Proses deteksi disini merupakan tahap implementasi program yang nantinya akan dibandingkan pada tahap pembahasan, baik SVM maupun CNN akan dibahas lebih rinci di tahap selanjutnya. Proses deteksi yang dimaksud adalah dimulai dari input data, *splitting data*, *training*, validasi, hingga *testing*. Hasil akhir yang diharapkan adalah program dapat mendeteksi objek kapal dengan menggunakan frame atau *bounding box* pada citra satelit beresolusi tinggi.
- f. Hasil dari kedua metode baik SVM dan CNN yang didapatkan dari tahap implementasi akan dibahas dan dibandingkan dengan menggunakan beberapa parameter tertentu.
- g. Menyimpulkan hasil penelitian dengan berbagai pertimbangan dan variable-variabel terkait.

### 3.2.2 Support vector machine

Model SVM dibangun menggunakan *scikit-image* dalam versi Python 3.7. Kami menggunakan fungsi *hog* *scikit-image* untuk mengekstrak fitur HOG. Terlepas dari fitur HOG, histogram warna dan fitur warna mentah juga digunakan. SVM dilatih dengan *train / val split* 3.200 / 800, ukuran gambar 80 x 80. Langkah-langkah pendeteksian kapal menggunakan metode SVM seperti yang ditunjukkan pada Gambar 3. 4.



Gambar 3. 4 *Flowchart* program dengan metode SVM

Penjelasan *flowchart* program dengan metode SVM:

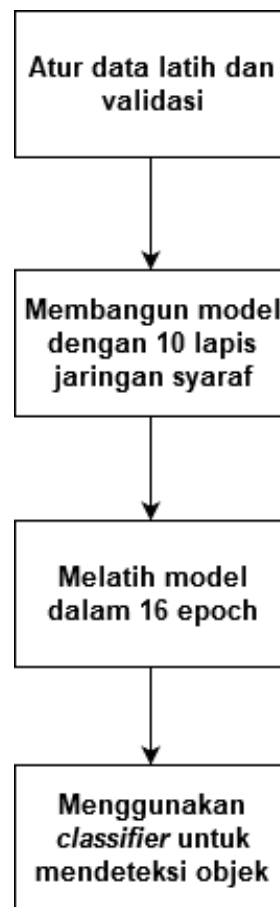
- h. Atur data latih dan validasi: pada proses ini data yang berisi 4.000 gambar akan di-split menjadi *training* dan *validation*. Pembagiannya berupa 80% atau 3.200 gambar untuk *training* dan 20% atau 800 gambar untuk *validation*.
- i. Ekstrak Ffitur citra: pada proses ekstraksi fitur digunakan tiga descriptor fitur yaitu *Histogram Of Gradients* (HOG), histogram warna, dan warna mentah. HOG, atau

*Histogram of Oriented Gradients*, adalah deskriptor fitur yang digunakan untuk mengekstrak fitur dari data gambar. Deskriptor HOG berfokus pada struktur atau bentuk objek (Singh, 2019). Untuk fitur color dan histogram color ukuran range atau bins yang digunakan adalah 32.

- j. Melatih *Classifier*: pada proses ini model dilatih dengan pengetahuan berupa data yang sudah di-split di awal yaitu sejumlah 3200 gambar.
- k. Menggunakan *classifier* untuk mendeteksi objek: Program mampu mendeteksi objek kapal dengan *sliding window* yang bergeser dengan jarak memberikan *bounding box* pada gambar yang diprediksi sebagai kapal.

### 3.2.3 Convolution neural networks

Kami membangun model CNN menggunakan Keras dari TensorFlow 2 dalam versi Python 3.7. CNN dilatih dengan *train / val split* 3.200 / 800, ukuran gambar 80 x 80. Langkah-langkah pendeteksian kapal menggunakan CNN dapat dilihat seperti pada Gambar 3. 5.



Gambar 3. 5 Flowchart program dengan metode CNN

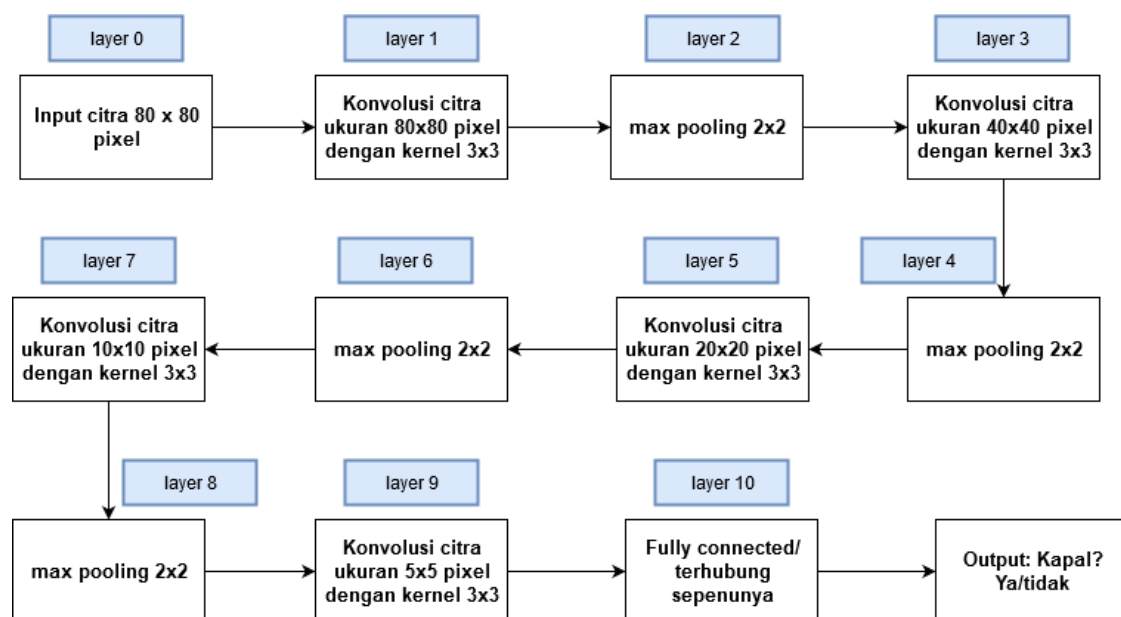
Penjelasan *flowchart* program dengan metode CNN:

- a. Atur data latih dan validasi: pada proses ini data yang berisi 4.000 gambar akan di-split menjadi *training* dan *validation*. Pembagiannya berupa 80% atau 3.200 gambar untuk *training* dan 20% atau 800 gambar untuk *validation*.
- b. Membangun model: Jaringan pada penelitian ini terdiri dari 10 lapisan CNN seperti yang ditunjukkan pada **Error! Reference source not found.**. Dalam jaringan ini, da proses konvolusi, max pooling, dan fully connecting.

- c. Melatih model: pada proses ini model dilatih dengan pengetahuan berupa data yang sudah di-split di awal yaitu sejumlah 3200 gambar dalam 16 kali iterasi.
- d. Menggunakan *classifier* untuk mendeteksi objek: Program mampu mendeteksi objek kapal dengan *sliding window* yang bergeser dengan jarak memberikan *bounding box* pada gambar yang diprediksi sebagai kapal.

Jaringan pada penelitian ini terdiri dari 10 lapisan CNN seperti yang ditunjukkan pada Gambar 3. 6. Dalam jaringan ini, ada proses konvolusi, max pooling, dan fully connecting. Proses konvolusi diterapkan tanpa *padding*, sehingga tidak mengubah ukuran gambar sebelum dan sesudah proses konvolusi. Tujuan max-pooling adalah untuk mengambil sampel yang mewakili input, mengurangi dimensinya dan memungkinkan untuk membuat asumsi tentang fitur yang terkandung dalam sub-wilayah yang terkubur. *Fully connected layer* menghubungkan setiap neuron dalam satu lapisan ke setiap neuron di lapisan lain.

Untuk proses *training* model akan dicoba dengan beberapa *epoch* yang berbeda. Pada penelitian ini model akan dijalankan dengan 6 *epoch* berbeda, yaitu *epoch* 15-20. Hasil dari ke-enam running program tersebut akan dibandingkan. Untuk proses deteksi pada citra satelit beresolusi tinggi akan digunakan *sliding window* yang bergerak atau bergeser dengan jarak 10 pixel pada setiap pergeseran. Ukuran *window* 80x80 disamakan dengan ukuran data *training*.



Gambar 3. 6 Susunan jaringan syaraf pada CNN